

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 21:09:07 ; Search time 156 Seconds
(without alignments)
11.498 Million cell updates/sec

Title: US-10-786-774-16
Perfect score: 28
Sequence: 1 DKCLA 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	28	100.0	5	6	AAG79805	Aag79805 SE motif
2	28	100.0	53	4	AAM94940	Aam94940 Human rep
3	28	100.0	81	4	ABB68303	Abb68303 Drosophil
4	28	100.0	220	6	ABU35339	Abu35339 Protein e
5	28	100.0	220	8	ADL05476	Adl05476 M. catarr
6	28	100.0	281	6	ABM67020	Abm67020 Phototrab
7	28	100.0	281	6	ABU25034	Abu25034 Protein e
8	28	100.0	331	4	ABB71974	Abb71974 Drosophil
9	28	100.0	363	3	AAY51660	Aay51660 Human AC1

10	28	100.0	363	3	AAY52031	Aay52031	Human	AC1
11	28	100.0	363	7	ADD14115	Add14115	Human	src
12	28	100.0	363	7	ADE58939	Ade58939	Human	Pro
13	28	100.0	363	7	ADE58943	Ade58943	Human	Pro
14	28	100.0	363	7	ADE58951	Ade58951	Human	Pro
15	28	100.0	363	7	ADE58947	Ade58947	Human	Pro
16	28	100.0	363	8	ADO19418	Ado19418	Human	PRO
17	28	100.0	364	7	ADE58949	Ade58949	Rat	Prote
18	28	100.0	364	7	ADE58945	Ade58945	Rat	Prote
19	28	100.0	364	7	ADE58941	Ade58941	Rat	Prote
20	28	100.0	364	7	ADE58937	Ade58937	Rat	Prote
21	28	100.0	445	7	ADF59424	Adf59424	Human	pol
22	28	100.0	574	4	AAG67370	Aag67370	Amino	aci
23	28	100.0	574	4	AAM52322	Aam52322	WASP	homo
24	28	100.0	660	8	ADM48103	Adm48103	Polypepti	
25	28	100.0	715	3	AAB19000	Aab19000	A	microtu
26	28	100.0	1477	2	AAY19981	Aay19981	B.	burgdo
27	28	100.0	1494	2	AAY19980	Aay19980	B.	burgdo
28	28	100.0	1494	6	ABU19204	Abu19204	Protein	e
29	26	92.9	28	8	ABO57504	Abo57504	Human	gen
30	26	92.9	67	5	ABB55054	Abb55054	Lactococc	
31	26	92.9	75	5	ABP29218	Abp29218	Streptoco	
32	26	92.9	82	6	ADA36306	Ada36306	Acinetoba	
33	26	92.9	83	5	ABP01140	Abp01140	Human	ORF
34	26	92.9	115	5	ABB49101	Abb49101	Listeria	
35	26	92.9	151	4	AAU17310	Aau17310	Novel	sig
36	26	92.9	151	7	ADB94018	Adb94018	Human	nov
37	26	92.9	215	5	ABB47460	Abb47460	Listeria	
38	26	92.9	238	2	AAY30349	Aay30349	Human	sur
39	26	92.9	238	6	ABO53004	Abo53004	Human	spl
40	26	92.9	238	7	ADB88953	Adb88953	Human	SMN
41	26	92.9	238	8	ADN49172	Adn49172	Human	sur
42	26	92.9	271	5	ABP39772	Abp39772	Staphyloc	
43	26	92.9	271	5	ABP41583	Abp41583	Human	ova
44	26	92.9	395	6	ABU20819	Abu20819	Protein	e
45	26	92.9	403	7	ABO82376	Abo82376	Pseudomon	

ALIGNMENTS

RESULT 1

AAG79805

ID AAG79805 standard; peptide; 5 AA.

XX

AC AAG79805;

XX

DT 16-APR-2003 (first entry)

XX

DE SE motif based peptide #3.

XX

KW Motif; SE; signal transduction; Alzheimer's disease; neuronal cell;

KW calreticulin; arrhythmia; idiopathic nephritic syndrome;

KW non-autoimmune hyperthyroidism; obesity; polycystic kidney disease;

KW cancer; asthma; atopic dermatitis; psoriasis; rheumatoid arthritis;

KW systemic lupus erythematosus; infectious disease; atherosclerosis;

KW inflammatory bowel disease; osteoarthritis; septic shock;

KW congestive heart failure; insulin-resistance syndrome;
 KW ischaemia-reperfusion injury.
 XX
 OS Homo sapiens.
 XX
 PN WO200299061-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US017536.
 XX
 PR 04-JUN-2001; 2001US-0295691P.
 PR 03-JUN-2002; 2002US-00161959.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Holoshitz J, Ling S;
 XX
 DR WPI; 2003-156853/15.
 XX
 PT Treating diseases associated with signal transduction aberrations, e.g.
 PT Alzheimer's disease, arrhythmia or rheumatoid arthritis, comprises
 PT administering a preparation comprising an SE- or an SE motif-containing
 PT peptide.
 XX
 PS Example 4; Page 63; 97pp; English.
 XX
 CC The sequences given in AAG79903-05 represent peptides which are based on
 CC the shared epitope (SE)-motif. These peptides were tested in DNA repair
 CC assys. Peptides containing the SE motifs may be used in the method of the
 CC invention for treating diseases associated with signal transduction
 CC aberrations. The method comprises: (a) providing a subject with one or
 CC more symptoms of Alzheimer's disease and, optionally, a plurality of
 CC neuronal cells expressing calreticulin, and a preparation comprising an
 CC SE-containing peptide or a peptide which binds the calreticulin; and (b)
 CC administering the preparation to the subject. The inventive method is
 CC useful for treating diseases associated with signal transduction
 CC aberrations, such as Alzheimer's disease, arrhythmia, idiopathic
 CC nephritic syndrome, non-autoimmune hyperthyroidism, obesity, polycystic
 CC kidney disease, cancer, asthma, atopic dermatitis, psoriasis, rheumatoid
 CC arthritis, systemic lupus erythematosus, infectious diseases,
 CC inflammatory bowel disease, osteoarthritis, septic shock,
 CC atherosclerosis, congestive heart failure, insulin-resistance syndrome,
 CC and ischaemia-reperfusion injury
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 6; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 1 DKCLA 5

RESULT 2

AAM94940

ID AAM94940 standard; protein; 53 AA.

XX

AC AAM94940;

XX

DT 21-NOV-2001 (first entry)

XX

DE Human reproductive system related antigen SEQ ID NO: 3598.

XX

KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.

XX

OS Homo sapiens.

XX

PN WO200155320-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001339.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Barash SC, Ruben SM;

XX

DR WPI; 2001-465570/50.

DR N-PSDB; AAL00910.

XX

PT Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.

XX

PS Claim 11; SEQ ID NO 3598; 1297pp + Sequence Listing; English.

XX

CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used

CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a protein of the invention
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 28; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 24 DKCLA 28

RESULT 3

ABB68303

ID ABB68303 standard; protein; 81 AA.

XX

AC ABB68303;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 31701.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR N-PSDB; ABL12406.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX

PS Disclosure; SEQ ID NO 31701, 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 81 AA;

Query Match 100.0%; Score 28; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 27 DKCLA 31

RESULT 4

ABU35339

ID ABU35339 standard; protein; 220 AA.

XX

AC ABU35339;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #20866.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Moraxella catarrhalis.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI; 2003-029926/02.

DR N-PSDB; ACA39209.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 25; SEQ ID NO 63263; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 220 AA;

Query Match 100.0%; Score 28; DB 6; Length 220;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 208 DKCLA 212

RESULT 5

ADL05476

ID ADL05476 standard; protein; 220 AA.

XX

AC ADL05476;

XX

DT 06-MAY-2004 (first entry)

XX

DE *M. catarrhalis* protein #1242.

XX

KW *Moraxella catarrhalis*; infection.

XX

OS *Moraxella catarrhalis*.

XX

PN US6673910-B1.

XX
 PD 06-JAN-2004.
 XX
 PF 04-APR-2000; 2000US-00540236.
 XX
 PR 08-APR-1999; 99US-0128416P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL;
 XX
 DR WPI; 2004-178127/17.
 DR N-PSDB; ADL03556.
 XX
 PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
 PT preparing a composition for diagnosing, preventing or treating infection
 PT caused by Moraxella catarrhalis.
 XX
 PS Disclosure; SEQ ID NO 3162; 429pp; English.
 XX
 CC The invention relates to an isolated nucleic acid encoding an Moraxella
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
 CC composition for diagnosing, preventing or treating infection caused by
 CC Moraxella catarrhalis. The present sequence represents the amino acid
 CC sequence of a M. catarrhalis protein.
 XX
 SQ Sequence 220 AA;

Query Match 100.0%; Score 28; DB 8; Length 220;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 208 DKCLA 212

RESULT 6

ABM67020

ID ABM67020 standard; protein; 281 AA.

XX

AC ABM67020;

XX

DT 20-NOV-2003 (first entry)

XX

DE Photorhabdus luminescens protein sequence #117.

XX

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.

XX

OS Photorhabdus luminescens.

XX

PN WO200294867-A2.

XX

PD 28-NOV-2002.

XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 2; SEQ ID NO 117; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 XX
 SQ Sequence 281 AA;

Query Match 100.0%; Score 28; DB 6; Length 281;
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 120 DKCLA 124

RESULT 7
 ABU25034
 ID ABU25034 standard; protein; 281 AA.
 XX
 AC ABU25034;
 XX

DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #10561.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Clostridium difficile.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA28904.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 52958; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 281 AA;

Query Match 100.0%; Score 28; DB 6; Length 281;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKCLA 5
| | | |
Db 37 DKCLA 41

RESULT 8

ABB71974

ID ABB71974 standard; protein; 331 AA.

XX

AC ABB71974;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 42714.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR N-PSDB; ABL16077.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX

PS Disclosure; SEQ ID NO 42714; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 331 AA;

Query Match 100.0%; Score 28; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
| | | | |
Db 229 DKCLA 233

RESULT 9

AAY51660

ID AAY51660 standard; protein; 363 AA.

XX

AC AAY51660;

XX

DT 01-JUN-2000 (first entry)

XX

DE Human AC12 protein fragment.

XX

KW Thermostable; template-dependent elongation; staple protein;

KW elongation protein; amplification; reverse transcription.

XX

OS Homo sapiens.

XX

PN WO200008164-A2.

XX

PD 17-FEB-2000.

XX

PF 06-AUG-1999; 99WO-DE002480.

XX

PR 06-AUG-1998; 98DE-01035653.

PR 07-SEP-1998; 98DE-01040771.

PR 18-JUN-1999; 99EP-00111795...

XX

PA (LION-) LION BIOSCIENCE AG.

XX

PI Kilger C, Kober I, Voss H, Moeckel G;

XX

DR WPI; 2000-195576/17.

XX

PT Thermostable in vitro polymerase complex for template-dependent

PT elongation of nucleic acids.

XX

PS Disclosure; Page 209-211; 233pp; German.

XX
CC This invention describes a novel thermostable in vitro complex (I) for
CC template-dependent elongation of nucleic acids (NA) comprising a
CC thermostable 'staple' protein and an elongation protein. The thermostable
CC in vitro complex is useful for template-dependent elongation of NA, e.g.
CC for amplification or reverse transcription. This is useful for sequencing
CC NA by the polymerase chain reaction or reverse transcriptase-PCR. The
CC complex can be used to mark NA. This sequence represents the human AC12
CC protein which is used to illustrate the method of the invention

XX
SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 3; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 335 DKCLA 339

RESULT 10
AAY52031

ID AAY52031 standard; protein; 363 AA.

XX

AC AAY52031;

XX

DT 12-JUL-2000 (first entry)

XX

DE Human AC12 protein.

XX

KW Polymerase; thermostable; nucleic acid elongation; sliding clamp protein;
KW amplification; reverse transcription.

XX

OS Homo sapiens.

XX

PN DE19840771-A1.

XX

PD 10-FEB-2000.

XX

PF 07-SEP-1998; 98DE-01040771.

XX

PR 06-AUG-1998; 98DE-01035653.

XX

PA (LION-) LION BIOSCIENCE AG.

XX

PI Voss H, Moeckel G, Kober I, Kilger C;

XX

DR WPI; 2000-195576/17.

XX

PT Thermostable in vitro polymerase complex for template-dependent
PT elongation of nucleic acids.

XX

PS Disclosure; Page 116-117; 152pp; German.

XX

CC This invention describes a novel thermostable in vitro complex for
CC template-dependent elongation of nucleic acids which comprises a

CC thermostable sliding clamp protein, which is connected with an elongation
CC protein that shows thermostable polymerase activity. The thermostable in
CC vitro accessory complex can be used to produce the thermostable in vitro
CC complex, which is useful for template-dependent elongation of nucleic
CC acids, e.g. for amplification or reverse transcription. This is useful
CC for sequencing nucleic acids by the polymerase chain reaction or reverse
CC transcriptase-polymerase chain reaction (RT-PCR). The complex can be used
CC to mark nucleic acids. AAY52000-Y52084 and AAW90752-W90799 represent
CC proteins and protein fragments used to illustrate the method of the
CC invention
XX
SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 3; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
| | | | |
Db 335 DKCLA 339

RESULT 11

ADD14115

ID ADD14115 standard; protein; 363 AA.

XX

AC ADD14115;

XX

DT 01-JAN-2004 (first entry)

XX

DE Human src biomarker polypeptide SEQ ID NO:304.

XX

KW predictor set; protein tyrosine kinase activity modulator;

KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;

KW gene therapy; drug sensitivity; genetic profile; cancer; human.

XX

OS Homo sapiens.

XX

PN WO2003062395-A2.

XX

PD 31-JUL-2003.

XX

PF 17-JAN-2003; 2003WO-US001981.

XX

PR 18-JAN-2002; 2002US-0350061P.

XX

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX

PI Huang F, Fairchild CR, Lee FY, Shaw P;

XX

DR WPI; 2003-636735/60.

DR N-PSDB; ADD14712.

XX

PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.

XX

PS Claim 10; SEQ ID NO 304; 139pp; English.

XX

CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.

XX

SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 7; Length 363;

Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5

||||

Db 335 DKCLA 339

RESULT 12

ADE58939

ID ADE58939 standard; protein; 363 AA.

XX

AC ADE58939;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein P35249, SEQ ID NO 4827.

XX

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; P35249.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 7; Length 363;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 335 DKCLA 339

RESULT 13

ADE58943

ID ADE58943 standard; protein; 363 AA.

XX

AC ADE58943;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein P35249, SEQ ID NO 4831.

XX

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

DR GENBANK; P35249.

XX

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 7; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 335 DKCLA 339

RESULT 14

ADE58951

ID ADE58951 standard; protein; 363 AA.

XX

AC ADE58951;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein P35249, SEQ ID NO 4839.

XX

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.
DR GENBANK; P35249.

XX

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 7; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKCLA 5
|||||
Db 335 DKCLA 339

RESULT 15

ADE58947

ID ADE58947 standard; protein; 363 AA.

XX

AC ADE58947;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein P35249, SEQ ID NO 4835.

XX

KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; P35249.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 363 AA;

Query Match 100.0%; Score 28; DB 7; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 335 DKCLA 339

Search completed: January 24, 2005, 21:23:42
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 21:20:28 ; Search time 39 Seconds
(without alignments)
8.502 Million cell updates/sec

Title: US-10-786-774-16
Perfect score: 28
Sequence: 1 DKCLA 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	28	100.0	73	4	US-09-621-976-5436	Sequence 5436, Ap
2	28	100.0	189	4	US-09-270-767-61188	Sequence 61188, A
3	28	100.0	220	4	US-09-540-236-3162	Sequence 3162, Ap
4	28	100.0	363	4	US-09-538-092-1072	Sequence 1072, Ap
5	28	100.0	447	4	US-09-270-767-45673	Sequence 45673, A
6	26	92.9	31	2	US-08-310-912A-88	Sequence 88, Appl
7	26	92.9	31	3	US-08-841-089-88	Sequence 88, Appl
8	26	92.9	31	3	US-09-301-085-88	Sequence 88, Appl
9	26	92.9	31	5	PCT-US95-04570-88	Sequence 88, Appl
10	26	92.9	31	5	PCT-US95-04589-88	Sequence 88, Appl
11	26	92.9	82	4	US-09-328-352-7593	Sequence 7593, Ap

12	26	92.9	223	4	US-09-270-767-39662	Sequence 39662, A
13	26	92.9	223	4	US-09-270-767-54879	Sequence 54879, A
14	26	92.9	238	3	US-09-028-327-1	Sequence 1, Appli
15	26	92.9	238	4	US-09-571-078A-1	Sequence 1, Appli
16	26	92.9	271	3	US-09-134-001C-4617	Sequence 4617, Ap
17	26	92.9	403	4	US-09-252-991A-31122	Sequence 31122, A
18	26	92.9	546	2	US-08-533-669A-2	Sequence 2, Appli
19	26	92.9	546	3	US-09-183-861-2	Sequence 2, Appli
20	26	92.9	546	3	US-09-022-765-2	Sequence 2, Appli
21	26	92.9	546	4	US-09-551-974A-2	Sequence 2, Appli
22	26	92.9	546	4	US-09-565-501A-2	Sequence 2, Appli
23	26	92.9	546	4	US-09-639-206A-2	Sequence 2, Appli
24	26	92.9	546	4	US-09-874-923-2	Sequence 2, Appli
25	26	92.9	546	4	US-08-798-841-2	Sequence 2, Appli
26	26	92.9	898	1	US-08-465-995A-2	Sequence 2, Appli
27	26	92.9	898	1	US-08-465-995A-4	Sequence 4, Appli
28	26	92.9	898	2	US-08-465-994C-2	Sequence 2, Appli
29	26	92.9	898	2	US-08-465-994C-4	Sequence 4, Appli
30	26	92.9	898	2	US-08-966-145-2	Sequence 2, Appli
31	26	92.9	898	2	US-08-966-145-4	Sequence 4, Appli
32	26	92.9	898	4	US-09-585-858-37	Sequence 37, Appl
33	26	92.9	920	1	US-08-101-593-2	Sequence 2, Appli
34	26	92.9	920	1	US-08-101-593-4	Sequence 4, Appli
35	26	92.9	982	4	US-09-551-974A-95	Sequence 95, Appl
36	26	92.9	982	4	US-09-565-501A-95	Sequence 95, Appl
37	26	92.9	982	4	US-09-639-206A-95	Sequence 95, Appl
38	26	92.9	982	4	US-09-874-923-95	Sequence 95, Appl
39	26	92.9	1427	4	US-09-551-974A-97	Sequence 97, Appl
40	26	92.9	1427	4	US-09-565-501A-97	Sequence 97, Appl
41	26	92.9	1427	4	US-09-639-206A-97	Sequence 97, Appl
42	26	92.9	1427	4	US-09-874-923-97	Sequence 97, Appl
43	26	92.9	1641	4	US-09-551-974A-96	Sequence 96, Appl
44	26	92.9	1641	4	US-09-565-501A-96	Sequence 96, Appl
45	26	92.9	1641	4	US-09-639-206A-96	Sequence 96, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-5436

; Sequence 5436, Application US/09621976

* Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 5436

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -49..-1
US-09-621-976-5436

Query Match 100.0%; Score 28; DB 4; Length 73;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 24 DKCLA 28

RESULT 2

US-09-270-767-61188

; Sequence 61188, Application US/09270767

*; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 61188

; LENGTH: 189

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-61188

Query Match 100.0%; Score 28; DB 4; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 107 DKCLA 111

RESULT 3

US-09-540-236-3162

; Sequence 3162, Application US/09540236

*; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 3162

; LENGTH: 220
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3162

Query Match 100.0%; Score 28; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 208 DKCLA 212

RESULT 4

US-09-538-092-1072

X; Sequence 1072, Application US/09538092
X; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1072
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P35249
US-09-538-092-1072

Query Match 100.0%; Score 28; DB 4; Length 363;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 335 DKCLA 339

RESULT 5

X US-09-270-767-45673
; Sequence 45673, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45673
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45673

Query Match 100.0%; Score 28; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
| | | |
Db 107 DKCLA 111

RESULT 6

US-08-310-912A-88

; Sequence 88, Application US/08310912A

; Patent No. 5981730

; GENERAL INFORMATION:

; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Staskawicz, Brian J.

; APPLICANT: Brent, Andrew F.

; APPLICANT: Dahlbeck, Douglas

; APPLICANT: Katagiri, Fumiaki

; APPLICANT: Kunkel, Barbara N.

; APPLICANT: Mindrinos, Michael N.

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION

; TITLE OF INVENTION: METHODS

; NUMBER OF SEQUENCES: 208

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2904

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/310,912A

; FILING DATE: September 22, 1994

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/227,360

; FILING DATE: April 13, 1994

```
; ATTORNEY/AGENT INFORMATION:
;   NAME: Lech, Karen F.
;   REGISTRATION NUMBER: 35,238
;   REFERENCE/DOCKET NUMBER: 00786/254001
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617) 542-5070
;   TELEFAX: (617) 542-8906
;   TELEX: 100254
; INFORMATION FOR SEQ ID NO: 88:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 31 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: not relevant
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-310-912A-88
```

```
Query Match          92.9%; Score 26; DB 2; Length 31;
Best Local Similarity 80.0%; Pred. No. 77;
Matches      4; Conservative 1; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 DKCLA 5
        |||:|
Db      21 DKCIA 25
```

RESULT 7

US-08-841-089-88

```
; Sequence 88, Application US/08841089
; Patent No. 6127607
```

; GENERAL INFORMATION:

```
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
```

```
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 106
```

; CORRESPONDENCE ADDRESS:

```
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2904
```

; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
```

; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/841,089
; FILING DATE:
; CLASSIFICATION:
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,360
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/230001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 100254
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-841-089-88

```

```

Query Match          92.9%; Score 26; DB 3; Length 31;
Best Local Similarity 80.0%; Pred. No. 77;
Matches      4; Conservative 1; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 DKCLA 5
        |||:|
Db      21 DKCIA 25

```

RESULT 8

```

US-09-301-085-88
; Sequence 88, Application US/09301085
; Patent No. 6262248
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; TITLE OF INVENTION: DETECTION METHODS
; FILE REFERENCE: 00786/254002
; CURRENT APPLICATION NUMBER: US/09/301,085
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: 08/310,912
; EARLIER FILING DATE: 1994-09-22
; EARLIER APPLICATION NUMBER: 08/227,360
; EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

```

US-09-301-085-88

Query Match 92.9%; Score 26; DB 3; Length 31;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKCLA 5
|||:
Db 21 DKCIA 25

RESULT 9

PCT-US95-04570-88

; Sequence 88, Application PC/TUS9504570

; GENERAL INFORMATION:

; APPLICANT: Ausubel, Frederick M.

; APPLICANT: Staskawicz, Brian J.

; APPLICANT: Brent, Andrew F.

; APPLICANT: Dahlbeck, Douglas

; APPLICANT: Katagiri, Fumiaki

; APPLICANT: Kunkel, Barbara N.

; APPLICANT: Mindrinos, Michael N.

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street Suite 3100

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2904

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04570

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/227,360

; FILING DATE: 13-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 00786/230001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: (617) 542-8906

; TELEX: 100254

; INFORMATION FOR SEQ ID NO: 88:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 31 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04570-88

Query Match 92.9%; Score 26; DB 5; Length 31;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 21 DKCIA 25

RESULT 10

PCT-US95-04589-88

; Sequence 88, Application PC/TUS9504589

; GENERAL INFORMATION:

; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 201
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2904

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04589
; FILING DATE:
; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/227,360
; FILING DATE: 13-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/230001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 100254

; INFORMATION FOR SEQ ID NO: 88:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids

; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04589-88

Query Match 92.9%; Score 26; DB 5; Length 31;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 21 DKCIA 25

RESULT 11

US-09-328-352-7593
; Sequence 7593, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7593
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7593

Query Match 92.9%; Score 26; DB 4; Length 82;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 22 DKCIA 26

RESULT 12

US-09-270-767-39662
; Sequence 39662, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39662
; LENGTH: 223

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39662

Query Match 92.9%; Score 26; DB 4; Length 223;
Best Local Similarity 80.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 195 DKCMA 199

RESULT 13

US-09-270-767-54879

; Sequence 54879, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54879
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54879

Query Match 92.9%; Score 26; DB 4; Length 223;
Best Local Similarity 80.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 195 DKCMA 199

RESULT 14

US-09-028-327-1

; Sequence 1, Application US/09028327
; Patent No. 6130064
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN SMN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.

```

; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,327
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0482 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT24
; CLONE: 3769729
US-09-028-327-1

```

```

Query Match          92.9%; Score 26; DB 3; Length 238;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches      4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 DKCLA 5
      |||:|
Db      77 DKCMA 81

```

RESULT 15

US-09-571-078A-1

```

; Sequence 1, Application US/09571078A
; Patent No. 6620783

```

GENERAL INFORMATION:

```

; APPLICANT: Tang, Y. Tom
; Corley, Neil C.
; Guegler, Karl J.
; TITLE OF INVENTION: HUMAN SMN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.

```

```

;      CITY: Palo Alto
;      STATE: CA
;      COUNTRY: USA
;      ZIP: 94304
;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette
;      COMPUTER: IBM Compatible
;      OPERATING SYSTEM: DOS
;      SOFTWARE: FastSEQ for Windows Version 2.0
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/571,078A
;      FILING DATE: 15-May-2000
;      CLASSIFICATION: <Unknown>
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: <Unknown>
;      FILING DATE: <Unknown>
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Billings, Lucy J.
;      REGISTRATION NUMBER: 36,749
;      REFERENCE/DOCKET NUMBER: PF-0482 US
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 650-855-0555
;      TELEFAX: 650-845-4166
;      TELEX: <Unknown>
;
;      INFORMATION FOR SEQ ID NO: 1:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 238 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;
;      IMMEDIATE SOURCE:
;      LIBRARY: BRSTNOT24
;      CLONE: 3769729
;
;      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-571-078A-1

```

```

Query Match          92.9%;  Score 26;  DB 4;  Length 238;
Best Local Similarity 80.0%;  Pred. No. 5.1e+02;
Matches      4;  Conservative    1;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 DKCLA 5
        |||:|
Db      77 DKCMA 81

```

```

Search completed: January.24, 2005, 21:28:25
Job time : 40 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 21:12:28 ; Search time 37 Seconds
(without alignments)
13.002 Million cell updates/sec

Title: US-10-786-774-16
Perfect score: 28
Sequence: 1 DKCLA 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	28	100.0	261	2	T32858	hypothetical prote	
2	28	100.0	281	1	I40679	beta-hydroxybutyry	
3	28	100.0	296	2	A45645	alpha-2-giardin -	
4	28	100.0	363	2	A45253	activator 1 37K ch	
5	28	100.0	367	2	S48833	cytochrome-c3 hydr	
6	28	100.0	553	2	AF2662	two component sens	
7	28	100.0	555	2	E97444	probable two-compo	
8	28	100.0	574	2	T43556	Wiskott-Aldrich sy	
9	28	100.0	574	2	T38819	wiskott-aldrich sy	
10	28	100.0	664	2	T33379	hypothetical prote	
11	28	100.0	1086	2	T18294	Ca2+-transporting	
12	28	100.0	1494	2	C70152	sensory transducti	
13	26	92.9	67	2	A86838	ferredoxin [import	

14	26	92.9	115	2	AC1134	hypothetical prote
15	26	92.9	150	2	H70164	hypothetical prote
16	26	92.9	167	2	FENTB	ferredoxin 2[4Fe-4
17	26	92.9	176	2	S00636	iron-sulfur protei
18	26	92.9	178	2	FERZB	ferredoxin 2[4Fe-4
19	26	92.9	180	2	S58619	ferredoxin 2[4Fe-4
20	26	92.9	183	2	FELVB	ferredoxin 2[4Fe-4
21	26	92.9	193	2	FEYBQI	NADH2 dehydrogenas
22	26	92.9	194	2	AH1834	NADH dehydrogenase
23	26	92.9	194	2	JQ2136	ferredoxin 2[4Fe-4
24	26	92.9	195	2	AG3592	zinc-finger protei
25	26	92.9	215	2	AD1171	transcription regu
26	26	92.9	251	2	A69897	hypothetical prote
27	26	92.9	591	2	S73790	hypothetical prote
28	26	92.9	725	2	A86328	protein F18014.27
29	26	92.9	818	2	F97772	hypothetical prote
30	26	92.9	898	1	DJBPT4	DNA-directed DNA p
31	26	92.9	1748	2	S63127	probable membrane
32	26	92.9	2493	2	S72349	nonstructural poly
33	26	92.9	2493	2	S26372	nonstructural poly
34	25	89.3	51	2	A91003	hypothetical prote
35	25	89.3	57	2	B38180	hypothetical prote
36	25	89.3	69	1	S48690	ubiquinol-cytochro
37	25	89.3	136	2	T25840	hypothetical prote
38	25	89.3	169	2	A87667	DNA invertase, pro
39	25	89.3	175	2	T26604	hypothetical prote
40	25	89.3	184	2	E87619	hypothetical prote
41	25	89.3	198	2	AB2835	hypothetical prote
42	25	89.3	198	2	E97612	hypothetical prote
43	25	89.3	207	2	AE0548	conserved hypothet
44	25	89.3	222	2	G64765	yaiV protein - Esc
45	25	89.3	222	2	A90682	hypothetical prote

ALIGNMENTS

RESULT 1

T32858

hypothetical protein T20H9.4 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32858

R;Blanchard, M.; Twyman, B.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of *C. elegans* cosmid T20H9.

A;Reference number: Z21235

A;Accession: T32858

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-261 <BLA>

A;Cross-references: UNIPROT:O44890; EMBL:AF040657; PIDN:AAB95053.1;

GSPDB:GN00021; CESP:T20H9.4

A;Experimental source: strain Bristol N2; clone T20H9

C;Genetics:

A;Gene: CESP:T20H9.4

A;Map position: 3

A;Introns: 5/3; 212/2

Query Match 100.0%; Score 28; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 125 DKCLA 129

RESULT 2

I40679

beta-hydroxybutyryl-CoA dehydrogenase, NAD-dependent (EC 1.1.1.-) - Clostridium difficile

C;Species: Clostridium difficile

C;Date: 12-Aug-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004

C;Accession: I40679; S49138

R;Mullany, P.; Clayton, C.L.; Pallen, M.J.; Slone, R.; al-Saleh, A.; Tabaqchali, S.

FEMS Microbiol. Lett. 124, 61-67, 1994

A;Title: Genes encoding homologues of three consecutive enzymes in the butyrate/butanol-producing pathway of Clostridium acetobutylicum are clustered on the Clostridium difficile chromosome.

A;Reference number: I40678; MUID:95095030; PMID:8001771

A;Accession: I40679

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-281 <RES>

A;Cross-references: UNIPROT:P45364; EMBL:X79899; NID:g509743; PIDN:CAA56272.1; PID:g509744

C;Genetics:

A;Gene: hbd

C;Superfamily: 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase homology

C;Keywords: NAD; oxidoreductase

F;1-280/Domain: 3-hydroxyacyl-CoA dehydrogenase homology <HCD>

F;2-30/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 100.0%; Score 28; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 37 DKCLA 41

RESULT 3

A45645

alpha-2-giardin - Giardia lamblia

N;Alternate names: giardin alpha chain 2

C;Species: Giardia lamblia

C;Date: 22-Apr-1993 #sequence_revision 24-May-1996 #text_change 09-Jul-2004

C;Accession: A45645

R;Alonso, R.A.; Peattie, D.A.

Mol. Biochem. Parasitol. 50, 95-104, 1992

A;Title: Nucleotide sequence of a second alpha giardin gene and molecular analysis of the alpha giardin genes and transcripts in Giardia lamblia.
A;Reference number: A45645; MUID:92178294; PMID:1542319
A;Accession: A45645
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-296 <ALO>
A;Cross-references: UNIPROT:P19389; GB:M34550; NID:g159099; PID:g159100
C;Genetics:
A;Introns: #status absent
A;Note: single copy gene
C;Superfamily: annexin I; annexin repeat homology
C;Keywords: cytoskeleton
F;25-74/Domain: annexin repeat homology #status atypical <AX1>

Query Match 100.0%; Score 28; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 257 DKCLA 261

RESULT 4

A45253

activator 1 37K chain - human

N;Alternate names: replication factor C 37K chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Aug-1998

C;Accession: A45253; B45253

R;Chen, M.; Pan, Z.Q.; Hurwitz, J.

Proc. Natl. Acad. Sci. U.S.A. 89, 5211-5215, 1992

A;Title: Studies of the cloned 37-kDa subunit of activator 1 (replication factor C) of HeLa cells.

A;Reference number: A45253; MUID:92302215; PMID:1351677

A;Accession: A45253

A;Molecule type: mRNA

A;Residues: 1-363 <CHE>

A;Experimental source: HeLa cells

A;Note: sequence extracted from NCBI backbone (NCBIN:106894, NCBIP:106895)

A;Accession: B45253

A;Molecule type: protein

A;Residues: 65-83;94-101;125-134;147-163;261-272 <CH2>

C;Superfamily: phage T4 DNA polymerase accessory protein 44

Query Match 100.0%; Score 28; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 335 DKCLA 339

RESULT 5

S48833

cytochrome-c3 hydrogenase (EC 1.12.2.1) alpha chain - *Pyrococcus furiosus*
 C;Species: *Pyrococcus furiosus*
 C;Date: 10-Dec-1994 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C;Accession: S48833
 R;Pedroni, P.
 submitted to the EMBL Data Library, September 1993
 A;Reference number: S48833
 A;Accession: S48833
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-367 <PED>
 A;Cross-references: UNIPROT:Q59667; EMBL:X75255; NID:g562773; PID:g563905
 C;Genetics:
 A;Start codon: GTG
 C;Superfamily: Cytochrome-c3 hydrogenase (EC 1.12.2.1)
 C;Keywords: oxidoreductase

Query Match 100.0%; Score 28; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 231 DKCLA 235

RESULT 6

AF2662

two component sensor kinase Atu0700 [imported] - *Agrobacterium tumefaciens*
 (strain C58, Dupont)

C;Species: *Agrobacterium tumefaciens*

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AF2662

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitajima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tingey, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.

A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF2662

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-553 <KUR>

A;Cross-references: UNIPROT:Q8UHI2; GB:AE008688; PIDN:AAL41716.1; PID:g17739064; GSPDB:GN00186

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu0700

A;Map position: circular chromosome

Query Match 100.0%; Score 28; DB 2; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 516 DKCLA 520

RESULT 7

E97444

probable two-component sensor (PA3974) [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)

C;Species: *Agrobacterium tumefaciens*

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: E97444

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, B.S.; Cao, Y.; Askenazi, M.; Halling, C.; Mullin, L.; Houmiel, K.; Gordon, J.; Vaudin, M.; Iartchouk, O.; Epp, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Flanagan, C.; Crowell, C.; Gurson, J.; Lomo, C.; Sear, C.; Strub, G.; Cielo, C.; Slater, S.
Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* C58.

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: E97444

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-555 <KUR>

A;Cross-references: UNIPROT:Q8UHI2; GB:AE007869; PIDN:AAK86510.1; PID:g15155664; GSPDB:GN00169

C;Genetics:

A;Gene: AGR_C_1264

A;Map position: circular chromosome

Query Match 100.0%; Score 28; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 518 DKCLA 522

RESULT 8

T43556

Wiskott-Aldrich syndrome protein homolog - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C;Accession: T43556

R;Zankel, T.C.; Ow, D.W.

submitted to the EMBL Data Library, December 1997

A;Description: A Wiskott-Aldrich Syndrome protein homolog in *Schizosaccharomyces pombe*, Wsplp, is implicated in stress-response pathways and control of the actin cytoskeleton.

A;Reference number: Z22575

A;Accession: T43556
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-574 <ZAN>
A;Cross-references: UNIPROT:O36027; EMBL:AF038575; PIDN:AAB92587.1
A;Experimental source: strain JS21
C;Genetics:
A;Gene: wsp1
A;Map position: I
A;Introns: 72/3; 519/3; 564/1

Query Match 100.0%; Score 28; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 102 DKCLA 106

RESULT 9

T38819

wiskott-aldrich syndrome protein homolog 1 - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T38819

R;Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997

A;Reference number: Z21813

A;Accession: T38819

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-574 <CON>

A;Cross-references: UNIPROT:O36027; EMBL:Z98980; NID:e1060691; PIDN:CAB11718.1;

GSPDB:GN00066; SPDB:SPAC4F10.15c

A;Experimental source: strain 972h-; cosmid c4F10

C;Genetics:

A;Gene: wsp1; SPDB:SPAC4F10.15c

A;Map position: 1

A;Introns: 72/3; 519/3; 564/1

Query Match 100.0%; Score 28; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 102 DKCLA 106

RESULT 10

T33379

hypothetical protein T07H3.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T33379

R;Beck, C.; O'Brien, D.; Kramer, J.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of *C. elegans* cosmid T07H3.
A;Reference number: Z21332
A;Accession: T33379
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-664 <BEC>
A;Cross-references: UNIPROT:O76612; EMBL:AF077540; PIDN:AAC26310.1;
GSPDB:GN00020; CESP:T07H3.1
A;Experimental source: strain Bristol N2; clone T07H3
C;Genetics:
A;Gene: CESP:T07H3.1
A;Map position: 2
A;Introns: 32/3; 79/1; 129/3; 309/3; 569/3; 601/1

Query Match 100.0%; Score 28; DB 2; Length 664;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 264 DKCLA 268

RESULT 11

T18294

Ca2+-transporting ATPase (EC 3.6.3.8) - *Entamoeba histolytica*

C;Species: *Entamoeba histolytica*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18294

R;Shen, P.; Yu, Y.; Samuelson, J.C.

submitted to the EMBL Data Library, January 1995

A;Description: Primary structure of a putative plasma membrane calcium ion-transporting ATPase.

A;Reference number: Z18863

A;Accession: T18294

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1086 <SHE>

A;Cross-references: UNIPROT:Q27642; EMBL:U20321; NID:g3392884; PID:g3392885;
PIDN:AAC28745.1

C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain homology

C;Keywords: ATP; hydrolase

F;629-800/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 100.0%; Score 28; DB 2; Length 1086;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 125 DKCLA 129

RESULT 12

C70152

sensory transduction histidine kinase homolog - Lyme disease spirochete

C;Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C;Accession: C70152

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman, J.; Utterback, T.; Watthey, L.; McDonald, L.; Artiach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: C70152

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1494 <KLE>

A;Cross-references: UNIPROT:O51381; GB:AE001146; GB:AE000783; NID:g2688312;

PIDN:AAC66780.1; PID:g2688313; TIGR:BB0420

A;Experimental source: strain B31

C;Keywords: phosphoprotein

F;1158-1275/Domain: response regulator homology <RRH>

F;1209/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 100.0%; Score 28; DB 2; Length 1494;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 1251 DKCLA 1255

RESULT 13

A86838

ferredoxin [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C;Species: *Lactococcus lactis* subsp. *lactis*

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C;Accession: A86838

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, S.D.; Sorokin, A.

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403.

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: A86838

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-67 <STO>

A;Cross-references: UNIPROT:Q9CEX9; GB:AE005176; PID:g12724722; PIDN:AAK05803.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: fer

Query Match 92.9%; Score 26; DB 2; Length 67;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 8 DKCIA 12

RESULT 14

AC1134

hypothetical protein lmo0474 [imported] - *Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C;Accession: AC1134

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1134

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-115 <GLA>

A;Cross-references: UNIPROT:Q8Y9Q0; GB:NC_003210; PIDN:CAC98553.1;

PID:g16409850; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo0474

Query Match 92.9%; Score 26; DB 2; Length 115;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||:|
Db 48..DKCMA 52

RESULT 15

H70164

hypothetical protein BB0521 - Lyme disease spirochete

C;Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C;Accession: H70164

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Kerlavage,

A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman, J.; Utterback, T.; Watthey, L.; McDonald, L.; Artiach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: H70164

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-150 <KLE>

A;Cross-references: UNIPROT:O51471; GB:AE001154; GB:AE000783; NID:g2688431;

PIDN:AAC66895.1; PID:g2688446; TIGR:BB0521

A;Experimental source: strain B31

Query Match 92.9%; Score 26; DB 2; Length 150;

Best Local Similarity 80.0%; Pred. No. 2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5

|||:|

Db 16 DKCIA 20

Search completed: January 24, 2005, 21:27:42

Job time : 39 secs

OM protein - protein search, using sw model

Run on: January 24, 2005, 21:27:09 ; Search time 143 Seconds
(without alignments)
12.632 Million cell updates/sec

Title: US-10-786-774-16
Perfect score: 28
Sequence: 1 DKCLA 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28	100.0	5	14	US-10-161-959-16	Sequence 16, Appl
2	28	100.0	5	17	US-10-786-774-16	Sequence 16, Appl
3	28	100.0	51	16	US-10-437-963-134248	Sequence 134248,
4	28	100.0	53	10	US-09-764-891-3598	Sequence 3598, Ap
5	28	100.0	91	17	US-10-425-115-344329	Sequence 344329,
6	28	100.0	92	16	US-10-437-963-168700	Sequence 168700,
7	28	100.0	114	16	US-10-767-701-39662	Sequence 39662, A
8	28	100.0	129	14	US-10-156-761-7804	Sequence 7804, Ap
9	28	100.0	147	17	US-10-425-115-225764	Sequence 225764,
10	28	100.0	150	16	US-10-767-701-33432	Sequence 33432, A
11	28	100.0	151	15	US-10-424-599-163019	Sequence 163019,
12	28	100.0	182	17	US-10-425-115-309204	Sequence 309204,
13	28	100.0	196	16	US-10-437-963-108022	Sequence 108022,
14	28	100.0	220	15	US-10-282-122A-63263	Sequence 63263, A
15	28	100.0	253	16	US-10-767-701-39663	Sequence 39663, A
16	28	100.0	278	17	US-10-425-115-309202	Sequence 309202,
17	28	100.0	281	15	US-10-282-122A-52958	Sequence 52958, A
18	28	100.0	309	17	US-10-425-115-309205	Sequence 309205,
19	28	100.0	310	15	US-10-424-599-233585	Sequence 233585,
20	28	100.0	524	16	US-10-767-701-45171	Sequence 45171, A
21	28	100.0	542	17	US-10-425-115-225767	Sequence 225767,
22	28	100.0	574	14	US-10-168-097A-76	Sequence 76, Appl
23	28	100.0	574	14	US-10-239-431A-38	Sequence 38, Appl
24	28	100.0	639	14	US-10-369-493-2275	Sequence 2275, Ap
25	28	100.0	660	14	US-10-310-154-521	Sequence 521, App
26	28	100.0	689	16	US-10-437-963-135053	Sequence 135053,
27	28	100.0	715	9	US-09-431-226-2	Sequence 2, Appli
28	28	100.0	1160	16	US-10-437-963-120406	Sequence 120406,
29	28	100.0	1287	16	US-10-437-963-120405	Sequence 120405,
30	28	100.0	1494	15	US-10-282-122A-47128	Sequence 47128, A
31	26	92.9	28	14	US-10-029-386-31138	Sequence 31138, A
32	26	92.9	31	9	US-09-867-852-88	Sequence 88, Appl
33	26	92.9	31	15	US-10-613-472-88	Sequence 88, Appl
34	26	92.9	31	16	US-10-613-765-88	Sequence 88, Appl
35	26	92.9	77	17	US-10-425-115-332185	Sequence 332185,
36	26	92.9	96	16	US-10-437-963-125954	Sequence 125954,
37	26	92.9	120	17	US-10-425-115-214445	Sequence 214445,
38	26	92.9	138	16	US-10-767-701-61312	Sequence 61312, A
39	26	92.9	151	9	US-09-764-868-875	Sequence 875, App
40	26	92.9	178	15	US-10-424-599-284096	Sequence 284096,
41	26	92.9	180	16	US-10-437-963-170304	Sequence 170304,
42	26	92.9	180	16	US-10-767-701-40967	Sequence 40967, A
43	26	92.9	191	15	US-10-425-114-43663	Sequence 43663, A
44	26	92.9	204	16	US-10-767-701-36137	Sequence 36137, A
45	26	92.9	207	15	US-10-424-599-178960	Sequence 178960,

ALIGNMENTS

RESULT 1
 US-10-161-959-16
 ; Sequence 16, Application US/10161959
 ; Publication No. US20030096748A1

```
; GENERAL INFORMATION:
; APPLICANT: Holoshitz, Joseph
; APPLICANT: Ling, Song
; TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases
Associated with
; TITLE OF INVENTION: Signal Transduction Aberrations
; FILE REFERENCE: UM-07135
; CURRENT APPLICATION NUMBER: US/10/161,959
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,691
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-161-959-16
```

```
Query Match          100.0%; Score 28; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 DKCLA 5
        |||||
Db      1 DKCLA 5
```

```
RESULT 2
US-10-786-774-16
; Sequence 16, Application US/10786774
; Publication No. US20040236071A1
; GENERAL INFORMATION:
; APPLICANT: Holoshitz, Joseph
; APPLICANT: Ling, Song
; TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases
Associated
; TITLE OF INVENTION: with Signal Transduction Aberrations
; FILE REFERENCE: UM-08550
; CURRENT APPLICATION NUMBER: US/10/786,774
; CURRENT FILING DATE: 2004-02-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-786-774-16
```

```
Query Match          100.0%; Score 28; DB 17; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 DKCLA 5
|||||
Db 1 DKCLA 5

RESULT 3

US-10-437-963-134248
; Sequence 134248, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134248
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36040C.1.pep
US-10-437-963-134248

Query Match 100.0%; Score 28; DB 16; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 40 DKCLA 44

RESULT 4

US-09-764-891-3598
; Sequence 3598, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3598
; LENGTH: 53

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-891-3598

Query Match 100.0%; Score 28; DB 10; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 24 DKCLA 28

RESULT 5

US-10-425-115-344329

; Sequence 344329, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 344329

; LENGTH: 91

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(91)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_77189C.1.pep

US-10-425-115-344329

Query Match 100.0%; Score 28; DB 17; Length 91;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 64 DKCLA 68

RESULT 6

US-10-437-963-168700

; Sequence 168700, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168700
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6718C.1.pep
US-10-437-963-168700
```

```
Query Match          100.0%; Score 28; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy          1 DKCLA 5
            |||||
Db          71 DKCLA 75
```

RESULT 7

```
US-10-767-701-39662
; Sequence 39662, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39662
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2555_2.pep
US-10-767-701-39662
```

```
Query Match          100.0%; Score 28; DB 16; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 DKCLA 5
|||||
Db 29 DKCLA 33

RESULT 8

US-10-156-761-7804
; Sequence 7804, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7804
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7804

Query Match 100.0%; Score 28; DB 14; Length 129;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 88 DKCLA 92

RESULT 9

US-10-425-115-225764
; Sequence 225764, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 225764
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_137489C.1.pep
US-10-425-115-225764

Query Match 100.0%; Score 28; DB 17; Length 147;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 135 DKCLA 139

RESULT 10

US-10-767-701-33432
; Sequence 33432, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 33432
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C31344_1.pep
US-10-767-701-33432

Query Match 100.0%; Score 28; DB 16; Length 150;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 91 DKCLA 95

RESULT 11

US-10-424-599-163019
; Sequence 163019, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163019
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118223C.1.pep
US-10-424-599-163019

Query Match 100.0%; Score 28; DB 15; Length 151;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 69 DKCLA 73

RESULT 12

US-10-425-115-309204

; Sequence 309204, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 309204
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(182)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_45060C.1.pep
US-10-425-115-309204

Query Match 100.0%; Score 28; DB 17; Length 182;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 42 DKCLA 46

RESULT 13

US-10-437-963-108022

; Sequence 108022, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 108022
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12316C.1.pep
US-10-437-963-108022

Query Match 100.0%; Score 28; DB 16; Length 196;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 129 DKCLA 133

RESULT 14

US-10-282-122A-63263

; Sequence 63263, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

```

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63263
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-282-122A-63263

```

```

Query Match          100.0%; Score 28; DB 15; Length 220;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 DKCLA 5
        |||||
Db      208 DKCLA 212

```

RESULT 15

```

US-10-767-701-39663
; Sequence 39663, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701

```

; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39663
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2555_1.pep
US-10-767-701-39663

Query Match 100.0%; Score 28; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 90 DKCLA 94

Search completed: January 24, 2005, 21:39:26
Job time : 145 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 21:11:48 ; Search time 192 Seconds
(without alignments)
14.984 Million cell updates/sec

Title: US-10-786-774-16
Perfect score: 28
Sequence: 1 DKCLA 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	100.0	67	2	Q83DV7	Q83dv7 coxiella bu
2	28	100.0	71	2	Q8CMD9	Q8cmd9 corynebacte
3	28	100.0	81	2	Q9W1B2	Q9w1b2 drosophila
4	28	100.0	81	2	O57094	O57094 cowpox viru
5	28	100.0	96	2	Q72ER2	Q72er2 desulfovibr
6	28	100.0	96	2	AAS94989	Aas94989 desulfovi
7	28	100.0	110	2	Q82R90	Q82r90 streptomyce
8	28	100.0	151	2	Q71HN3	Q71hn3 ficus awkeo
9	28	100.0	151	2	AAQ07268	Aaq07268 ficus awk
10	28	100.0	261	2	O44890	O44890 caenorhabdi
11	28	100.0	280	1	END4_PHOLL	Q7n364 photorhabdu
12	28	100.0	281	1	HBD_CLODI	P45364 clostridium
13	28	100.0	296	1	GIA2_GIALA	P19389 giardia lam
14	28	100.0	296	2	Q7QXC9	Q7qxc9 giardia lam
15	28	100.0	315	2	Q6MX78	Q6mx78 azoarcus sp

16	28	100.0	315	2	CAF21980	Caf21980 azoarcus
17	28	100.0	326	2	Q8MQN1	Q8mqn1 drosophila
18	28	100.0	331	2	Q9VJ32	Q9vj32 drosophila
19	28	100.0	340	2	Q6TRS1	Q6trsl ustilago ma
20	28	100.0	340	2	AAQ94939	Aaq94939 ustilago
21	28	100.0	358	2	Q6DRF4	Q6drf4 brachydanio
22	28	100.0	360	2	Q729F1	Q729f1 desulfovibr
23	28	100.0	360	2	AAS96873	Aas96873 desulfovi
24	28	100.0	363	1	RFC4_HUMAN	P35249 homo sapien
25	28	100.0	363	2	AAP35633	Aap35633 homo sapi
26	28	100.0	364	2	Q99J62	Q99j62 mus musculu
27	28	100.0	367	2	Q59667	Q59667 pyrococcus
28	28	100.0	367	2	Q8U2E5	Q8u2e5 pyrococcus
29	28	100.0	391	2	Q6TBE1	Q6tbel dillenia ph
30	28	100.0	391	2	AAR30463	Aar30463 dillenia
31	28	100.0	414	2	Q6Q880	Q6q880 leptosphaer
32	28	100.0	414	2	AAS92548	Aas92548 leptospha
33	28	100.0	478	2	Q8A9R4	Q8a9r4 bacteroides
34	28	100.0	532	2	Q9CKT3	Q9ckt3 pasteurella
35	28	100.0	537	2	Q865A4	Q865a4 sus scrofa
36	28	100.0	553	2	Q98MF6	Q98mf6 rhizobium l
37	28	100.0	553	2	Q8UHI2	Q8uhi2 agrobacteri
38	28	100.0	555	2	Q7D0W5	Q7d0w5 agrobacteri
39	28	100.0	574	2	O36027	O36027 schizosacch
40	28	100.0	669	2	Q7Q7U7	Q7q7u7 anopheles g
41	28	100.0	683	2	O76612	O76612 caenorhabdi
42	28	100.0	715	2	Q9I8N0	Q9i8n0 xenopus lae
43	28	100.0	716	2	Q6NUF4	Q6nuf4 xenopus lae
44	28	100.0	716	2	AAH68637	Aah68637 xenopus l
45	28	100.0	718	2	Q7R2A7	Q7r2a7 giardia lam

ALIGNMENTS

RESULT 1

Q83DV7

ID Q83DV7 PRELIMINARY; PRT; 67 AA.
AC Q83DV7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CBU0585;
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;

RT "Complete genome sequence of the Q-fever pathogen, Coxiella
RT burnetii.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR EMBL; AE016961; AA090129.1; -.
DR TIGR; CBU0585; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 67 AA; 8006 MW; B59308FC612A568D CRC64;

Query Match 100.0%; Score 28; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 20 DKCLA 24

RESULT 2

Q8CMD9

ID Q8CMD9 PRELIMINARY; PRT; 71 AA.
AC Q8CMD9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created).
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CE2796, CE2806;
OS Corynebacterium efficiens YS-314.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=196164;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP005223; BAC19606.1; -.
DR EMBL; AP005223; BAC19616.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 71 AA; 8349 MW; 4ABF5B41FC7D0925 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 6 DKCLA 10

RESULT 3

Q9W1B2

ID Q9W1B2 PRELIMINARY; PRT; 81 AA.
 AC Q9W1B2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE CG13564-PA.
 GN ORFNames=CG13564;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,

RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AE003463; AAF47160.1; -.
 DR FlyBase; FBgn0034973; CG13564.
 DR InterPro; IPR007512; DUF543.
 DR Pfam; PF04418; DUF543; 1.
 SQ SEQUENCE 81 AA; 8817 MW; 44AEF7947EAFCD FDD CRC64;

Query Match 100.0%; Score 28; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 27 DKCLA 31

RESULT 4

O57094

ID O57094 PRELIMINARY; PRT; 81 AA.

AC O57094;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Tumor necrosis factor receptor II homolog.
 GN Name=crmD;
 OS Cowpox virus (CPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10243;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Munich OPV89/5;
 RX MEDLINE=98188282; PubMed=9520445;
 RA Loparev V.N., Parsons J.M., Knight J.C., Panus J.F., Ray C.A.,
 RA Buller R.M.L., Pickup D.J., Esposito J.J.;
 RT "A third distinct tumor necrosis factor receptor of orthopoxviruses."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3786-3791(1998).
 DR EMBL; U87582; AAB94352.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 1.
 DR PROSITE; PS00652; TNFR_NGFR_1; 1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 1.
 KW Receptor.
 SQ SEQUENCE 81 AA; 9160 MW; DA80737087B9214D CRC64;

Query Match 100.0%; Score 28; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKCLA 5
 |||||
 Db 65 DKCLA 69

RESULT 5.

Q72ER2

ID Q72ER2 PRELIMINARY; PRT; 96 AA.
 AC Q72ER2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=DVU0507;
 OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
 OS 8303).
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
 OC Desulfovibrionaceae; Desulfovibrio.
 OX NCBI_TaxID=882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15077118; DOI=10.1038/nbt959;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,

RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT *Desulfovibrio vulgaris* Hildenborough.";
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AE017310; AAS94989.1; -.
 DR TIGR; DVU0507; -.
 DR InterPro; IPR007546; DUF503.
 DR Pfam; PF04456; DUF503; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 96 AA; 10779 MW; B7CE58B47F0B2140 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 71 DKCLA 75

RESULT 6

AAS94989

ID AAS94989 PRELIMINARY; PRT; 96 AA.
 AC AAS94989;
 DT 27-APR-2004 (TrEMBLrel. 27, Created)
 DT 27-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN DVU0507.
 OS *Desulfovibrio vulgaris* (strain Hildenborough / ATCC 29579 / NCIMB
 OS 8303).
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
 OC Desulfovibrionaceae; Desulfovibrio.
 OX NCBI_TaxID=882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15077118;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT *Desulfovibrio vulgaris* Hildenborough.";
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AE017310; AAS94989.1; -.
 DR TIGR; DVU0507; -.
 KW Hypothetical protein.
 SQ SEQUENCE 96 AA; 10779 MW; B7CE58B47F0B2140 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 71 DKCLA 75

RESULT 7

Q82R90

ID Q82R90 PRELIMINARY; PRT; 110 AA.
AC Q82R90;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MutT-family protein.
GN OrderedLocusNames=SAV253;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AP005021; BAC67962.1; -.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR Pfam; PF00293; NUDIX; 1.
KW Complete proteome.
SQ SEQUENCE 110 AA; 12355 MW; 80EA7047366DB53C CRC64;

Query Match 100.0%; Score 28; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 69 DKCLA 73

RESULT 8

Q71HN3

ID Q71HN3 PRELIMINARY; PRT; 151 AA.

AC Q71HN3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Ripening induced protein.
 OS Ficus awkeotsang (Jelly fig).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Rosales; Moraceae; Ficus.
 OX NCBI_TaxID=204231;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Maturing achene;
 RA Hsu J.S.F., Tzen J.T.C.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF497748; AAQ07268.1; -.
 DR InterPro; IPR000916; Bet_v_I.
 DR Pfam; PF00407; Bet_v_I; 1.
 DR ProDom; PD000531; Bet_v_I; 1.
 SQ SEQUENCE 151 AA; 17151 MW; FE9B7E468CE9DCE7 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 151;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 110 DKCLA 114

RESULT 9

AAQ07268

ID AAQ07268 PRELIMINARY; PRT; 151 AA.
 AC AAQ07268;
 DT 10-MAY-2004 (TrEMBLrel. 27, Created)
 DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Ripening induced protein.
 OS Ficus awkeotsang.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Rosales; Moraceae; Ficus.
 OX NCBI_TaxID=204231;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Maturing achene;
 RA Hsu J.S.F., Tzen J.T.C.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF497748; AAQ07268.1; -.
 SQ SEQUENCE 151 AA; 17151 MW; FE9B7E468CE9DCE7 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 151;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||

Db 110 DKCLA 114

RESULT 10

O44890

ID O44890 PRELIMINARY; PRT; 261 AA.
AC O44890;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein T20H9.4.
GN ORFNames=T20H9.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Blanchard M., Twyman B.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF040657; AAB95053.1; -.
DR PIR; T32858; T32858.
DR WormPep; T20H9.4; CE13862.
DR InterPro; IPR002900; DUF38.
DR Pfam; PF01827; FTH; 1.
KW Hypothetical protein.
SQ SEQUENCE 261 AA; 30649 MW; 8832E3BC175A02DF CRC64;

Query Match 100.0%; Score 28; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKCLA 5
|||||
Db 125 DKCLA 129

RESULT 11

END4_PHOLL

ID END4_PHOLL STANDARD; PRT; 280 AA.
AC Q7N364;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Probable endonuclease IV (EC 3.1.21.2) (Endodeoxyribonuclease IV).
 GN Name=nfo; OrderedLocusNames=plu2857;
 OS Photorhabdus luminescens (subsp. laumondii).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Photorhabdus.
 OX NCBI_TaxID=141679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TT01;
 RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
 RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
 RA Dassa E., Deroose R., Derzelle S., Freyssinet G., Gaudriault S.,
 RA Medigue C., Lanois A., Powell K., Siguier P., Vincent R., Wingate V.,
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
 RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
 RT luminescens.";
 RL Nat. Biotechnol. 21:1307-1313(2003).
 CC -!- FUNCTION: Endonuclease IV plays a role in DNA repair. It cleaves
 CC phosphodiester bonds at apurinic or apyrimidinic sites (AP sites)
 CC to produce new 5' ends that are base-free deoxyribose 5-phosphate
 CC residues. It preferentially attacks modified AP sites created by
 CC bleomycin and neocarzinostatin (By similarity).
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphooligonucleotide end-products.
 CC -!- COFACTOR: Binds 3 zinc ions (By similarity).
 CC -!- SIMILARITY: Belongs to the AP endonuclease family 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; BX571868; CAE15231.1; -.
 DR PhotoList; plu2857; -.
 DR HAMAP; MF_00152; -; 1.
 DR InterPro; IPR001719; AP_endnuclease2.
 DR Pfam; PF01261; AP_endonuc_2; 1.
 DR TIGRFAMs; TIGR00587; nfo; 1.
 DR PROSITE; PS00729; AP_NUCLEASE_F2_1; 1.
 DR PROSITE; PS00730; AP_NUCLEASE_F2_2; 1.
 DR PROSITE; PS00731; AP_NUCLEASE_F2_3; FALSE_NEG.
 KW Complete proteome; DNA repair; Endonuclease; Hydrolase; Metal-binding;
 KW Nuclease; Zinc.
 FT METAL 69 69 Zinc 1 (By similarity).
 FT METAL 109 109 Zinc 1 (By similarity).
 FT METAL 145 145 Zinc 1 and 2 (By similarity).
 FT METAL 179 179 Zinc 2 (By similarity).
 FT METAL 182 182 Zinc 3 (By similarity).
 FT METAL 216 216 Zinc 2 (By similarity).
 FT METAL 229 229 Zinc 3 (By similarity).
 FT METAL 231 231 Zinc 3 (By similarity).
 FT METAL 261 261 Zinc 2 (By similarity).

SQ SEQUENCE 280 AA; 31389 MW; 444818E39B91480B CRC64;

Query Match 100.0%; Score 28; DB 1; Length 280;

Best Local Similarity 100.0%; Pred. No. 5.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DKCLA 5
|||||
Db 120 DKCLA 124

RESULT 12

HBD_CLODI

ID HBD_CLODI STANDARD; PRT; 281 AA.

AC P45364;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (Beta-

DE hydroxybutyryl-CoA dehydrogenase) (BHBD).

GN Name=hbd;

OS Clostridium difficile.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1496;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=E;

RX MEDLINE=95095030; PubMed=8001771;

RA Mullany P., Clayton C.L., Pallen M.J., Slone R., Al-Saleh A.,

RA Tabaqchali S.;

RT "Genes encoding homologues of three consecutive enzymes in the
RT butyrate/butanol-producing pathway of Clostridium acetobutylicum are
RT clustered on the Clostridium difficile chromosome.";

RL FEMS Microbiol. Lett. 124:61-67(1994).

CC -!- CATALYTIC ACTIVITY: (S)-3-hydroxybutanoyl-CoA + NADP(+) = 3-
CC acetoacetyl-CoA + NADPH.

CC -!- PATHWAY: Butyrate/butanol-producing pathway.

CC -!- SIMILARITY: Belongs to the 3-hydroxyacyl-CoA dehydrogenase family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X79899; CAA56272.1; -.

DR PIR; I40679; I40679.

DR HSSP; P00348; 3HDH.

DR InterPro; IPR006180; 3HCDH.

DR InterPro; IPR006108; 3HCDH_C.

DR InterPro; IPR006176; 3HCDH_N.

DR InterPro; IPR008927; 6DGDH_C_like.

DR Pfam; PF00725; 3HCDH; 1.

DR Pfam; PF02737; 3HCDH_N; 1.

DR PROSITE; PS00067; 3HCDH; 1.
KW Fatty acid metabolism; NADP; Oxidoreductase.
SQ SEQUENCE 281 AA; 30668 MW; 192357C9B050E98C CRC64;

Query Match 100.0%; Score 28; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 37 DKCLA 41

RESULT 13

GIA2_GIALA

ID GIA2_GIALA STANDARD; PRT; 296 AA.
AC P19389;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Giardin alpha-2 chain.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Portland-1;
RX MEDLINE=92178294; PubMed=1542319;
RA Alonso R.A., Peattie D.A.;
RT "Nucleotide sequence of a second alpha giardin gene and molecular
RT analysis of the alpha giardin genes and transcripts in Giardia
RT lamblia.";
RL Mol. Biochem. Parasitol. 50:95-104(1992).
CC -!- FUNCTION: Giardins are involved in parasite attachment to the
CC intestinal mucosa and in the cytoskeletal disassembly and
CC reassembly that marks the transition from infectious trophozoite
CC to transmissible cyst. They may interact with other cytoskeletal
CC proteins such as microtubules in the microribbons or crossbridges,
CC to maintain the integrity of the ventral disk.
CC -!- SUBCELLULAR LOCATION: Most likely in the edges of the ventral disk
CC microribbons.
CC -!- SIMILARITY: Contains 1 annexin repeat.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34550; AAA29150.1; -.
DR PIR; A45645; A45645.
DR HSSP; P04083; 1B09.
DR InterPro; IPR008088; Alpha_giardin.
DR InterPro; IPR001464; Annexin.
DR Pfam; PF00191; Annexin; 2.

DR PRINTS; PR01712; ALPHAGIARDIN.
KW Annexin; Cytoskeleton; Microtubule; Multigene family.
FT REPEAT 4 69 Annexin.
SQ SEQUENCE 296 AA; 33935 MW; AA9C79312C716A54 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 257 DKCLA 261

RESULT 14

Q7QXC9

ID Q7QXC9 PRELIMINARY; PRT; 296 AA.
AC Q7QXC9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP_741_29684_28794.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000061; EAA39694.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; IEA.
DR InterPro; IPR008088; Alpha_giardin.
DR InterPro; IPR001464; Annexin.
DR Pfam; PF00191; Annexin; 2.
DR PRINTS; PR01712; ALPHAGIARDIN.
DR ProDom; PD000143; Annexin; 1.
SQ SEQUENCE 296 AA; 33935 MW; AA9C79312C716A54 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
|||||
Db 257 DKCLA 261

RESULT 15

Q6MX78

ID Q6MX78 PRELIMINARY; PRT; 315 AA.
 AC Q6MX78;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sugar dehydratase.
 GN Name=fn1299; ORFNames=c2A203;
 OS Azoarcus sp. (strain EbN1).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
 OC Rhodocyclaceae; Azoarcus.
 OX NCBI_TaxID=76114;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EbN1;
 RA Kube M., Heider J., Amann J., Hufnagel P., Kuehner S., Beck A.,
 RA Reinhardt R., Rabus R.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EbN1;
 RA PROSCIENCE;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BX682953; CAF21980.1; -.
 SQ SEQUENCE 315 AA; 35864 MW; E39B0F16B420B32B CRC64;

Query Match 100.0%; Score 28; DB 2; Length 315;
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DKCLA 5
 |||||
 Db 135 DKCLA 139

Search completed: January 24, 2005, 21:27:00
 Job time : 196 secs